

Special Section on HDR Imaging

Spatially varying image based lighting using HDR-video



Jonas Unger*, Joel Kronander, Per Larsson, Stefan Gustavson, Joakim Löw, Anders Ynnerman

Media and Information Technology, Linköping University, Sweden

ARTICLE INFO

Article history:

Received 1 March 2013

Received in revised form

4 July 2013

Accepted 9 July 2013

Available online 20 July 2013

Keywords:

High dynamic range video

Image based lighting

Scene capture and processing

Photo realistic rendering

ABSTRACT

Illumination is one of the key components in the creation of realistic renderings of scenes containing virtual objects. In this paper, we present a set of novel algorithms and data structures for visualization, processing and rendering with real world lighting conditions captured using High Dynamic Range (HDR) video. The presented algorithms enable rapid construction of general and editable representations of the lighting environment, as well as extraction and fitting of sampled reflectance to parametric BRDF models. For efficient representation and rendering of the sampled lighting environment function, we consider an adaptive (2D/4D) data structure for storage of light field data on proxy geometry describing the scene. To demonstrate the usefulness of the algorithms, they are presented in the context of a fully integrated framework for spatially varying image based lighting. We show reconstructions of example scenes and resulting production quality renderings of virtual furniture with spatially varying real world illumination including occlusions.

© 2013 The Authors. Published by Elsevier Ltd. Open access under CC BY-NC-SA license.

1. Introduction

The production of photo-realistic computer graphics renderings, seamlessly merging synthetic objects into real world scenes, is a fundamental goal in computer graphics. A difficult and time consuming challenge is to accurately model the synthetic lighting to match that of the real environment. This has motivated the development of *Image Based Lighting* (IBL) techniques [1], where the light in real world scenes is captured and used as a source of illumination in renderings. IBL has recently been extended from using a single HDR light probe measurement to include sequences of High Dynamic Range (HDR) images captured at different locations in the scene [2,3] capturing the full 5D spatial and angular dimensionality of the illumination. This generalization enables the production of highly realistic IBL renderings in environments exhibiting complex spatial variations with preserved parallax. However, these methods have previously been constrained to small scenes in controlled environments, and require extensive manual processing to produce good results. Furthermore previous methods are limited to simplistic representations which make post-processing and editing of the captured data difficult, and in some cases even impossible.

In this paper, we present a novel framework for building detailed representations of real world illumination captured in general scenes that can be edited, and re-used, and which produce superior image quality compared to previous methods. The framework is based on recent developments in HDR video capture [4,5] which enables efficient capture of the full dimensionality of the illumination in large environments exhibiting both complex spatial and angular variations. From the comparison in Fig. 1 between traditional IBL 1a and our method 1b, it is evident that the realism and visual interest in the resulting renderings increase significantly by the lighting complexity enabled by HDR video based scene capture.

The HDR video sequences are often tens or hundreds of GB in size and pose great challenges for efficient storage, processing, and rendering. We present solutions to these challenges, and integrate the algorithms into a functioning software toolset and corresponding data structures. Our framework consists of a number of coupled hardware and software components for HDR video based scene capture and reconstruction and modeling of both the geometric and the radiometric properties of the scene, as well as data structures for efficient rendering of the reconstructed scene illumination. Our pipeline can, as illustrated in Fig. 2, be divided into five steps:

1. *Scene capture*: The scene (indoor or outdoor) is captured using a combination of panoramic HDR light probe sequences and ordinary HDR video sequences with a smaller field of view.
2. *Scene reconstruction and modeling*: The goal of the scene reconstruction and modeling step is to generate a geometric proxy model of the scene, and to extract the light sources in the scene. This is carried out as a semi-automatic, purely

* Corresponding author. Tel.: +46 11 363436.
E-mail address: jonas.unger@liu.se (J. Unger).
URL: <http://vcl.itn.liu.se> (J. Unger).



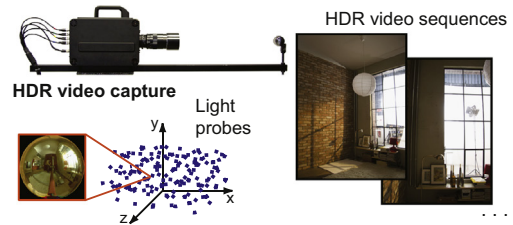
Fig. 1. A comparison between (a) traditional image based lighting using a single HDR light probe image, and (b) rendering with our framework generated from thousands of light probes captured using HDR video. By including the complex spatial and angular variations in the scene illumination the proposed technique enables a new level of realism compared to previous IBL techniques.

image-based process where we use a combination of structure from motion techniques and a novel interactive modeling approach based on shape from shading that we call *focal volume* modeling.

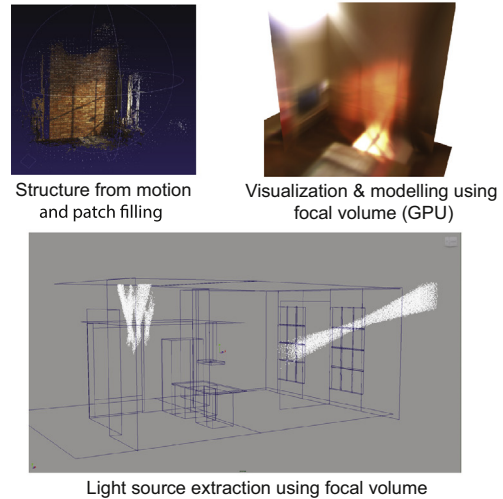
3. **Radiance re-projection:** The radiance data captured in the HDR sequences is then re-projected onto the proxy geometry where it is stored as view dependent textures (light fields) in an adaptive 2D/4D data structure which enables fast lookup during rendering.
4. **Light field processing:** All points on all objects have not been imaged from all directions, because the sample region is of finite size and resolution, and there may be occlusions in the scene. This leads to holes in the re-projected radiance data. The final step is to fill in such holes by the mean or interpolated value in the angular domain or, if the geometry is accurately recovered, by fitting a parametric BRDF to the captured data.
5. **Rendering:** The resulting model can be directly used for rendering of highly realistic images of virtual objects placed into the real scene, using a rendering approach that is straightforward to incorporate in most rendering packages.

To show the utility of the approach, we take our examples from IKEA Communications AB, the creators of the most widely distributed publication in the world—the IKEA Catalogue. We show how the presented framework can be applied to rapid generation of ultra realistic renderings of virtually furnished scenes.

1. Scene capture using HDR video

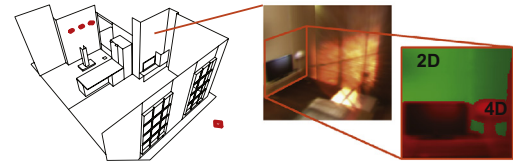


2. Scene reconstruction and modeling

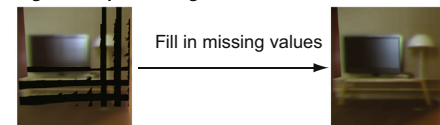


3. Radiance re-projection

The captured scene radiance data is re-projected and stored in an adaptive 2D/4D data structure



4. Light field processing



5. Rendering



Fig. 2. Our systems pipeline can be divided into five steps: (1) HDR video based scene capture, (2) scene reconstruction and modeling, where a geometric proxy model of the scene is built, (3) radiance re-projection, where the capture HDR data is projected onto the model, (4) a processing step where holes in the re-projected radiance data are filled in, followed by (5) placement of virtual objects and rendering. Each step is described in detail in the following sections.

2. Background

The central component in image synthesis is to estimate the outgoing radiance $L(\mathbf{x} \rightarrow \vec{\omega}_o)$ leaving a surface point \mathbf{x} in an outgoing direction $\vec{\omega}_o$. This is commonly modeled by the rendering

equation [6] as

$$L(\mathbf{x} \rightarrow \vec{\omega}_o) = \int_{\Omega} L(\mathbf{x} \leftarrow \vec{\omega}_i) f_r(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_o) \cos \theta_i d\vec{\omega}_i \quad (1)$$

where $L(\mathbf{x} \leftarrow \vec{\omega}_i)$ denotes the incident radiance distribution arriving at \mathbf{x} from direction $\vec{\omega}_i$, $f_r(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_o)$ is the surface BRDF describing the scattering of light from incident directions $\vec{\omega}_i$ to the outgoing direction $\vec{\omega}_o$, and $\cos \theta_i$ describes the projected area resulting from the angle, θ_i , between the surface normal at \mathbf{x} and the incident direction, $\vec{\omega}_i$.

Image based lighting: Traditional IBL [1] uses a single panoramic HDR image, a light probe, captured at a fixed point \mathbf{x}_c describing the scene illumination as an environment map. Each pixel in the light probe image can be thought of as the radiance incident onto the virtual objects over a small solid angle from the real scene. During rendering of a point \mathbf{x} in the scene as observed from an outgoing direction $\vec{\omega}_o$, traditional IBL approximates the rendering equation (1) as

$$L(\mathbf{x} \rightarrow \omega_o) = \int_{\Omega} \hat{L}_{\text{IBL}}(\mathbf{x}_c \leftarrow \vec{\omega}_i) f_r(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_o) V(\mathbf{x}, \vec{\omega}_i) \cos \theta_i d\omega_i \quad (2)$$

where $\hat{L}_{\text{IBL}}(\mathbf{x}_c \leftarrow \vec{\omega}_i)$ is the angular radiance distribution captured in the light probe at position \mathbf{x}_c , and $V(\mathbf{x}, \vec{\omega}_i)$ is the visibility function ($V=1$ if the captured environment map is visible in the direction $\vec{\omega}_i$ and $V=0$ if it is occluded by inserted virtual geometry). As $\hat{L}_{\text{IBL}}(\mathbf{x}_c \leftarrow \vec{\omega}_i)$ is independent of the surface position on the virtual object, \mathbf{x} , traditional IBL effectively treats the lighting as being identical across the entire scene and originating from infinitely far away.

The development of HDR imaging techniques has now made it possible to capture sequences of light probe images, for an overview see [7]. Inspired by light field and reflectance field imaging techniques [8–11], IBL has been extended to capture the 5D plenoptic function [12] including both spatial and angular variations in the scene illumination. Such radiance data sets are referred to as Incident Light Fields (ILF) [2,3,13,14]. For efficient ILF rendering, it is necessary to build a geometric proxy model of the scene onto which the captured 5D radiance data can be re-projected and stored as 4D light fields. For this purpose, unstructured ILF data has been resampled to planes [15], bounding boxes around the entire scene [16], and 3D grids of spherical harmonics inside the captured region [17].

The ILF methods render a surface point \mathbf{x} on a virtual object using an approximation of the rendering equation, (1), as

$$L(\mathbf{x} \rightarrow \omega_o) = \int_{\Omega} \hat{L}_{\text{SIBL}}(\mathbf{x} \leftarrow \vec{\omega}_i) f_r(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_o) V(\mathbf{x}, \vec{\omega}_i) \cos \theta_i d\omega_i \quad (3)$$

where $\hat{L}_{\text{SIBL}}(\mathbf{x} \leftarrow \vec{\omega}_i)$ is the interpolated incident illumination at point \mathbf{x} given measured data in the neighborhood of the point.

Previous techniques for spatially varying image based lighting have been constrained to small scenes [2,15] and controlled environments [3,16,17]. In this work, we extend these methods to fully general scenes by introducing an efficient workflow for scene capture, scene modeling and editing, and rendering. Our framework enables faster capture and produces significantly more detailed representations of the scene lighting than previous work, as well as more intuitive editing and interpolation of captured radiance data.

Advanced photo editing: Another related application area is to edit single photographs by rendering virtual objects into the captured scene [18] by building approximate lighting models based on information in the image, or editing the appearance of existing objects by exploiting flaws in the human visual system [19]. These are powerful techniques. They are, however, significantly different in that the goal is to edit a single photograph as compared to the proposed framework in this paper, where the

goal is to build a general and physically accurate model of the lighting environment that allows virtual objects to be placed anywhere in the scene and be rendered with photo-realistic results

Modeling scene geometry: Recent automatic scene reconstruction methods based on structure from motion techniques [20,21], multiple view stereo methods [22], and the use of additional range sensors such as in the Microsoft Kinect [23,24] provide high quality results. However, the reconstructed geometry does not necessarily reflect the artistic requirements and the desired level of detail, especially for general scenes with unknown depth complexity, non-planar geometry, textureless and non-Lambertian surfaces. Furthermore, our goal is not to extract detailed scene geometry, but to approximate light field surfaces for efficient storage and processing of radiance information. We therefore use a semi-automatic approach with user control that guarantees a predictable and robust result. In contrast to previous image based modeling techniques developed for a relatively small set of planar 8-bit input images [25–27], our approach is targeted at the handling of a very large number of HDR light probe images. Such HDR input data with a panoramic mapping is what enables practical, dense and efficient sampling of the 5D radiance distribution during scene capture [1,2,16]. Inspired by synthetic aperture imaging [28,29] we use focus as a visual cue to guide the user in the process of extracting objects in the scene. Extending this idea to use HDR input data allows for the extraction of light sources as 4D light fields.

3. Scene capture using HDR-video

Recent advances in imaging hardware and processing software have enabled the direct capture of HDR Video [30,4,31,5]. Using HDR video, several thousands of light probes can be captured in a few minutes. Compared to previous spatially varying IBL methods [2,16,15], this enables rapid, unconstrained capture of detailed lighting environments in general scenes.

To measure the incident illumination in the scene, both high resolution HDR video sequences and panoramic HDR light probe sequences are captured. The panoramic sequences ensure full angular coverage at each position within the capture region, enabling efficient and accurate measurement of the incident light field in the scene. The high resolution HDR video sequences are used for structure from motion synthesis of scene geometry and as backdrop images during rendering. The camera position and orientation can be estimated with sufficient accuracy using standard computer vision techniques [32]. For the panoramic light probes, however, this is often difficult in practice due to e.g. non-single view point projections as for the commonly used mirror sphere (see Fig. 3 for an example light probe setup). For the experiments in this paper, the position and orientation of the light probe sequences were instead measured using external tracking. For our experiments, we have used a custom built mechanical tracking stage capable of tracking the motion of the camera within a $1.5 \times 1.5 \times 1.5 \text{ m}^3$ cube with an accuracy of 0.1 mm, see Fig. 3d. For larger capture areas, the stage can be moved and the data sets aligned in the processing software. It is also possible to use high accuracy optical tracking based on external cameras and tracking markers.

The HDR video sequences used for the experiments presented in this paper were captured using two different camera setups. They are both designed to output raw uncompressed HDR frames at video frame rates. The first setup, described in detail in [30,4], is based on the camera platform Ranger C55 from the sensor and camera manufacturer SICK-IVP. The camera platform is an off-the-shelf

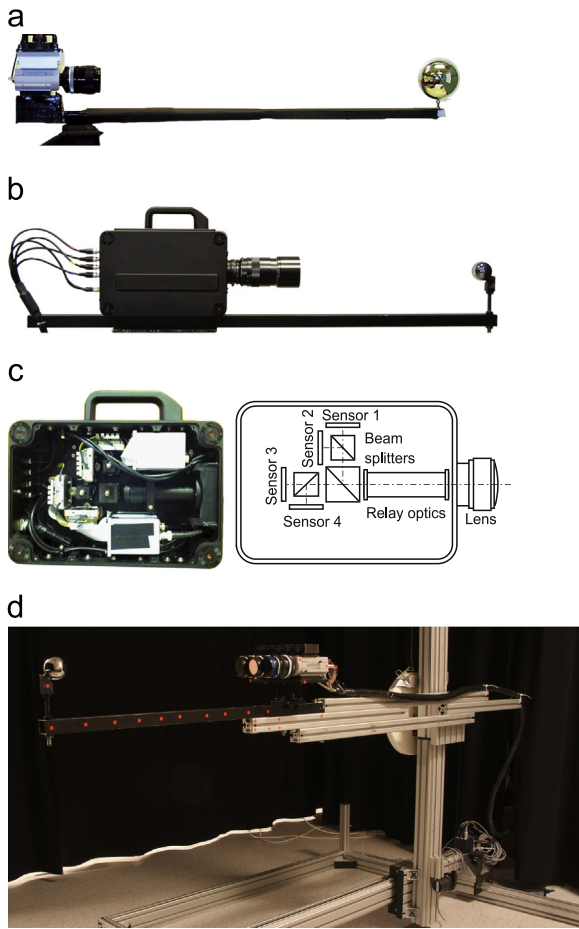


Fig. 3. Displays the HDR-video cameras used in the experiments in this paper. (a) shows the SICK-IVP Ranger C55 in a mirror sphere light probe setup. (b) the multi-sensor camera with global shutter captures HDR images with a resolution of 2336×1752 exhibiting a dynamic range of more than 24 f-stops. (c) displays the layout of the multi-sensor system. (d) displays an HDR-video setup mounted on a mechanical tracking stage, where the position of the mirror sphere is tracked within a $1.5 \times 1.5 \times 1.5 \text{ m}^3$ cube with an accuracy of 0.1 mm.

industrial inspection camera built around a “smart” CMOS sensor. The micro-code running on the sensor chip was modified to enable sequential HDR capture with minimized time disparity between the different exposures in a rolling shutter fashion. The setup exhibits a dynamic range of 23 f-stops at 25 fps with a resolution of 960×512 pixels. Fig. 3a shows the camera in a light probe setup. The second HDR camera setup, described in [5], is a custom built multi-sensor setup with a global shutter, see Fig. 3b. The HDR capture is based on four high quality Kodak KAI-04050 CCD sensors with a resolution of 2336×1752 pixels and RGB Bayer pattern CFA¹ sampling. The sensors image the scene through a common front lens, but the light path is split four ways by beam splitters and passed through four different ND filters, see Fig. 3c. The relay lens, which extends the optical path length to make room for the beam splitters, was custom designed for this setup, which is one of the keys to the high quality of the system. Each sensor has 12 bits linear A/D conversion, and the exposure levels of the four sensors cover a range of $1 : 2^{12}$, which yields a dynamic range equivalent to $12 + 12 = 24$ bits of linear resolution, or “24 f-stops”. The dynamic range can be extended further by varying the exposure times between the sensors.

The camera sensors are connected to a host computer through a CameraLink interface. This system allows for capture, processing and offline storage of up to 32 frames per second at 4 Mpixels resolution, amounting to around 1 GB/s of raw data.

It should be noted that the presented pipeline and the supporting algorithms do not rely on any specific camera hardware for HDR video capture, and both the above systems, although very different in their technical specification, enable rapid capture with little distortion and noise. The only requirement on the input data is that it captures the natural illumination in the scene of interest with a sufficiently large dynamic range such that the data can be trusted as accurate measurements of the scene radiance information. This excludes some commercially available HDR video systems as they are only capable of outputting tone mapped 8-bit data, which is not sufficient for accurate photometric measurements.

4. Scene reconstruction and modeling

During scene capture, large amounts of HDR video footage are captured. For moderate sized scenes the complete data set is often in the order of ≈ 100 –200 GB. However, the captured data is highly redundant, and in its raw form, it is a poor representation of the incident light in the scene. To enable efficient processing, rendering and editing of the captured radiance data, a geometric proxy model of the scene is reconstructed. This proxy geometry can then be used as a more efficient incident light field representation by back-projecting the captured radiance data to the modeled surfaces.

The goal of the modeling step is thus to build a dense geometric proxy model of the scene consisting of a set of 2D surfaces, A_i , onto which the illumination information in the sampled radiance data set, \mathcal{L} , can be projected and stored. The geometric modeling is carried out in a semi-automatic fashion where the user is kept in the loop to enable efficient and editable reconstructions. While a fully automatic geometric reconstruction would be possible, using e.g. recent methods such as Kinect Fusion [23,24], these reconstructions tend to be overly detailed and too noisy for the sole purpose of storing radiance information. Another advantage of using a simple representation of the scene geometry, consisting only of piecewise planar proxy geometries, is that it enables efficient editing of the captured scene illumination in a post-process.

4.1. Structure from motion

The first step in the modeling step is the reconstruction of a dense point cloud describing the scene using well-known structure from motion techniques. In our pipeline implementation, we use the publicly available frameworks Bundler [32] and PMVS2 [33] to obtain a set of dense 3D patches with normals describing the geometry of diffuse surfaces in the scene. An example point cloud is displayed in Fig. 4a. The point clouds generally approximate diffuse surfaces well, and it is possible to automatically generate a full polygonal model of the scene based on the point clouds. However, for transmissive and specular surfaces and highly complex scenes it is necessary to adjust and often also model parts of the scene by hand due to inaccuracies in the obtained models. To do this, we use an interactive approach based on shape from focus, presented in the next section.

4.2. Interactive modeling using focal volumes

In the second step, the user interactively builds a dense piecewise planar model of the scene based on the recovered point

¹ A Color Filter Array (CFA) is usually placed in front of the monochrome sensor for color imaging.

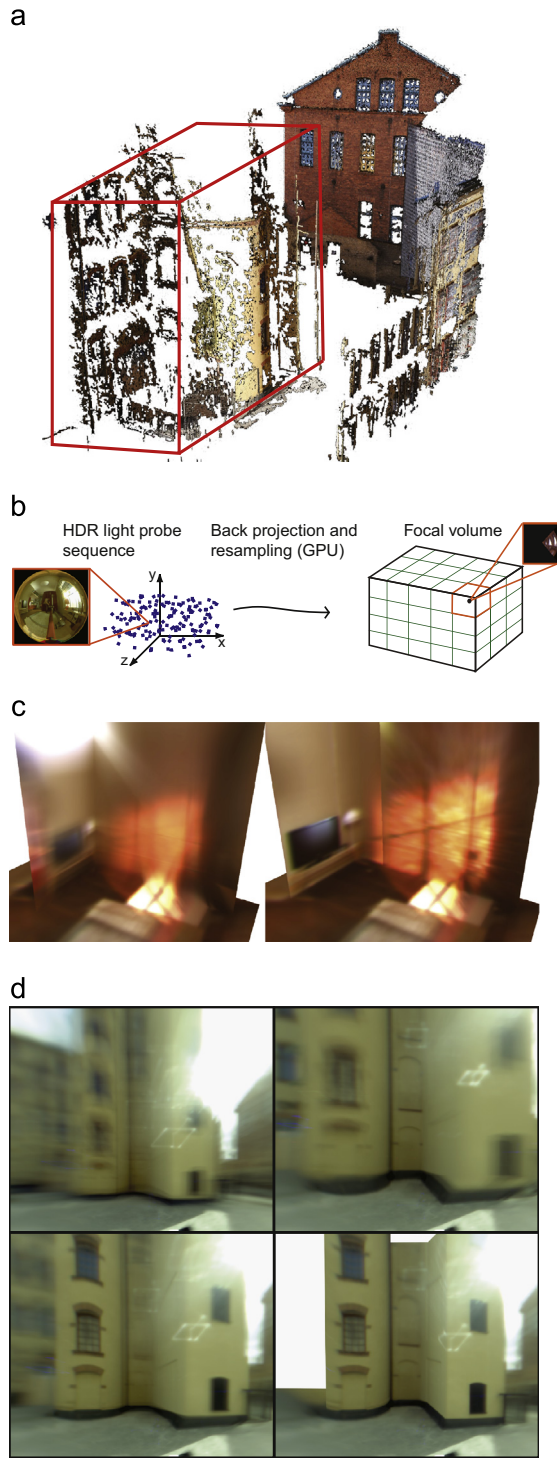


Fig. 4. (a) A point cloud recovered from an outdoor scene with a user specified focal volume (red). (b) Each pixel in each light probe image in the sequences is back projected from its capture position, \mathbf{x}_i , along its corresponding direction, \vec{d}_i , and contributes to the statistics computed at each voxel it intersects. (c) and (d) Exploiting focus as a strong visual cue, the focal volumes enable the user to rapidly model or modify the scene geometry. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

cloud and shape from focus cues. The modeling can be carried out using a commercial 3D modeling package such as Maya. To provide the user with an intuitive interface for modeling based on shape from focus, a set of 3D, voxel-based representations of the captured lighting environment is created. We refer to these data structures as *focal volumes*. Using this approach means that,

instead of working on the images directly as in traditional image based modeling [25–27], or on the captured full 5D data described by \mathcal{L} , we compute statistics of the measured scene radiance passing through each point in the scene and represent it as scalar-valued volumetric functions. This is related to the concept of radiant flux, although the flux only covers parts of the angular domain for points outside of the captured volume. Based on the point cloud, the user selects regions in space where focal volumes will be generated, see Fig. 4a. In order to reconstruct a scalar volume from the radiance samples \mathcal{L} , as illustrated in Fig. 4b, we use two different approaches. One is to use a splatting operation and re-project each sample, $l \in \mathcal{L}$, from its position, \mathbf{x}_i , on the mirror sphere backwards along its direction, \vec{d}_i , and deposit its radiance value at each intersected voxel. Another is to traverse all voxels and sample each probe image in the corresponding directions. The latter involves more work but enables better quality because filtering can be performed in the image domain during sampling, and it also parallelizes better. We have found that the statistics that are most intuitive to use are the radiance mean value, $\bar{w}(x, y, z)$, and its variance, $w_v(x, y, z)$, where the mean and variance are estimated based on the available angles of incidence. To account for the varying sample density in the mirror sphere images, and to remove unwanted artifacts such as the camera and the operator, we weight each radiance sample according to its position on the sphere. To preserve the dynamic range of the statistics, $\bar{w}(x, y, z)$ and $w_v(x, y, z)$, of the linear radiance samples without clipping, the focal volumes are stored and viewed as 32-bit or 16-bit floating point data.

Conceptually, a focal volume is a synthetic aperture projection using a 3D aperture of the size and shape of the capture region, and exhibits two key features that we exploit as visual cues to guide the user in the modeling process. A surface placed inside the volume can be thought of as a *focal surface*. Points on a focal surface that coincide with a surface in the real scene will be in focus, but will rapidly move out of focus with the distance to the real surface. Secondly, due to the HDR nature of the data, light sources will be represented by distinct high intensity voxels. Fig. 4c and d displays examples captured during interactive modeling of an outdoor and indoor scene, respectively. It should be noted that this is a truly interactive process and that it is best viewed in the supplementary video attached to this paper. Due to the large aperture, the synthetic depth of field is very narrow, and objects in the scene cause strong and robust local focus maxima.

In general we use around 1000 light probe images for the volume re-projection. However, if only a few light probes are used for the volume reconstruction, the defocus will instead manifest itself as misalignment of scene features. This is also a strong cue because of the Vernier acuity of the human visual system. The light field geometries, A_i , can be modeled interactively using these visual cues by sweeping surfaces and primitives in and out of focus. The synthetic aperture nature of the volume reveals partially occluded objects and resolves depth complexities. Heavily occluded parts of the volume can be re-traced with the occluding A_i surfaces as a visibility constraint. Using a subset of the samples in \mathcal{L} , this re-tracing can be performed incrementally at interactive speed during modeling. The final proxy geometries, A_i , for the example photo studio scene in Fig. 10a are presented in Fig. 5b.

4.3. Light source extraction

For efficient rendering, it is necessary to extract strong direct light sources in the scene and label them as light sources during rendering. In this way it is straightforward to treat them as direct illuminants or by other means sample them more densely than low intensity regions during rendering. As demonstrated in [16], this has the potential of decreasing rendering times by orders of

magnitude. As described in the previous section, the focal volumes are computed and stored as floating point data. This ensures that the full dynamic range of the input radiance samples is kept intact without clipping. Due to the HDR nature of the focal volumes, light sources will be represented by extremely intense voxels in the mean radiance volume, \bar{w} .

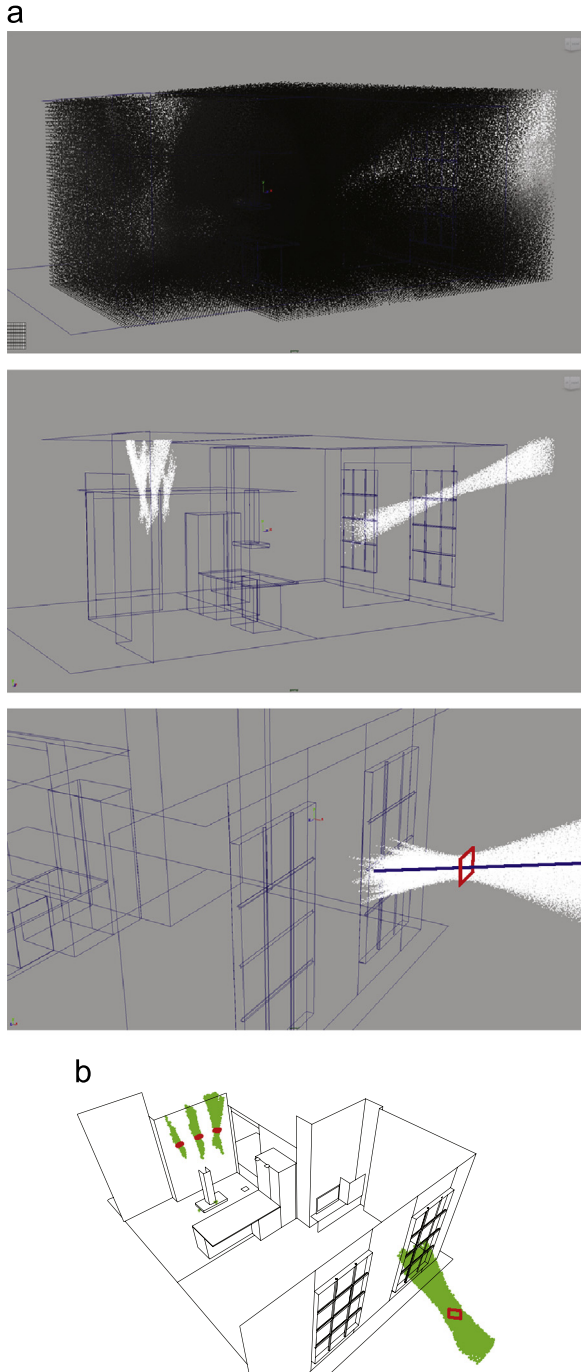


Fig. 5. (a) A focal volume interactively thresholded to reveal the direct light sources in the scene. Light sources are represented as cones of extremely intense voxels in the mean radiance volume. Each cone of high intensity voxels is first segmented using region growing from a user selected seed point. The largest principal component (blue line) of the cluster of voxels is found. The light source (red plane) is then found by fitting a plane through the cone apex, large enough to contain the cone points orthogonal to the largest principal component. (b) The final proxy geometry and extracted light sources from the real photo studio. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Light sources that are smaller than the sample region(s) will be characterized by cones of high intensity voxels. Fig. 5a displays how interactive thresholding of the focal volume robustly reveals the cones representing the light sources in the recovered scene model. This enables well-defined segmentation of the voxels corresponding to each light source, and makes it possible to extract its position and spatial extent using interactive spatial selection and thresholding. As the projected radiance samples converge, the voxel intensities increase to a distinct maximum at the position of the real light source. From a user-specified position on the cone and a threshold, the conic region is segmented using either region-growing or manual selection in the modeling software. The orientation of the cone is analyzed using principal component analysis. A plane, A_i , orthogonal to the strongest principal component is then constructed at the cone apex, covering the extent of the light source as specified by the user. For area light sources larger than the extent of the sample region(s), Γ , a cone of convergent rays cannot be found, but in this case it is possible to use the focus based modeling method described in the previous paragraph. Fig. 5b displays the extracted light sources (red) and their corresponding voxel cones for a high threshold setting (green).

5. Light field re-projection and data structure

The next step is to re-project the radiance samples captured in the HDR sequences to the surfaces A_i in the recovered geometric proxy model, where they are stored as 4D view dependent textures, light fields $L_{A_i}(u, v, \phi, \theta)$. Here (u, v) denotes the spatial coordinate on the A_i surface, and (ϕ, θ) coordinates in the angular domain.

To form the light fields, L_{A_i} , in the scene, the full set of captured radiance samples, \mathcal{L} , are re-projected to the proxy geometries, A_i , and stored in an adaptive 2D/4D data structure. This is performed either as direct sampling of the visible light probe images or by splatting each radiance sample, $l \in \mathcal{L}$, from its capture position, \mathbf{x}_l , along $\vec{\omega}_l$ to the first intersected surface, A_i . The decision on whether to store 2D or 4D data at a point is based on the local coefficient of variation, $c_v = \sqrt{w_v(x, y, z) / \bar{w}(x, y, z)}$. By visualizing $c_v(x, y, z)$, see Fig. 6, the user can specify a threshold to determine where the light field can be assumed to be diffuse and independent of angle (green). At such points L_{A_i} is compactly stored as a single value, saving significant amounts of data. At non-diffuse points (red), the angular variations from specularly or parallax are preserved and stored as small 2D angular maps.

For surfaces classified as diffuse, several candidate radiance samples $l_c \in \mathcal{L}$, from different capture positions, $\mathbf{x}_l \in \Gamma$ and directions $\vec{\omega}_l \in S^2$, could be used to describe the texture on the surface. Blending the candidate radiance samples, l_c , for each pixel results in ghosting when unmodeled geometric detail is present, and in general it blurs the textures when there are slight misalignments of the proxy geometry. Moreover, due to the spherical geometry used for capturing light probe images, the radiance samples, l_c , have a varying degree of accuracy depending on the imaged position on the mirror sphere. Strongly distorted samples are obtained close to the edge of the sphere, corresponding to forward-facing viewing angles, where considerable blurring and noise is introduced. If only the most accurate measured radiance value would be assigned to each single pixel, this would result in significant seams in the texture when regions of pixels are assigned to different images. To enforce smoothness constraints and at the same time maximize image quality, we use a Markov Random Field (MRF) optimization to obtain a high quality seamless texture map. Similar to previous work [34,20,35] we minimize an energy term consisting of data penalties and smoothness terms through a graph-cut optimization framework. In contrast to previous methods using images obtained

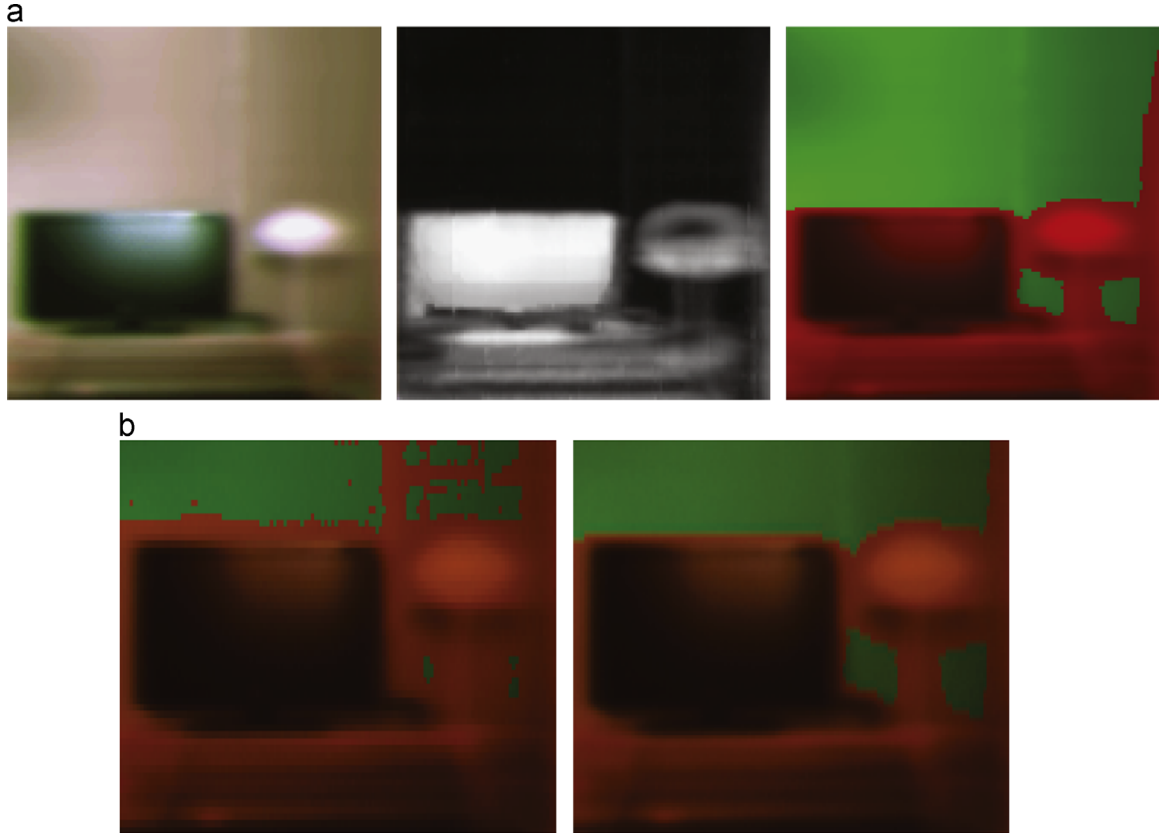


Fig. 6. (a) Interactive visualization of the (left) mean radiance, \bar{w} , (middle) radiance variance, w_v , and (right) user controlled 2D/4D thresholding. (b) Interactive visualization of the 2D/4D thresholding for user selection of how data is stored. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

without significant geometrical distortion, we also consider the accuracy with which the radiance samples, l_c , are captured, and use a smoothness term that favors similar spherical distortion of radiance samples used for neighboring pixel values.

The assignment of image values to pixels in the 2D texture of a light field surface, A_i , is carried out using a pixel labeling approach where each label b_p corresponds to a visible light probe image. To connect pixels and enforce smoothness constraints we use a 4-connected neighborhood consisting of vertical and horizontal neighbors. To find the optimal texturing we consider the following energy function:

$$E(b_p) = \sum_p D_p(b_p) + \sum_{p,q} V_{p,q}(b_p, b_q) \quad (4)$$

where p and q represent pixels in the 2D texture A_i , $D_p(b_p)$ is a data term that describes the cost of assigning the label b_p to pixel p and $V_{p,q}(b_p, b_q)$ describes the cost of assigning neighbor pixels p, q labels b_p and b_q . To solve the labeling problem, the energy function described by Eq. (4) is minimized using the max flow algorithm [36–38]. The data term is composed of a photo-consistency term, a preference for a frontal view and a term enforcing high spatial resolution of the samples. The photo-consistency term penalizes radiance samples deviating from the median radiance value over all samples, assuring a robust texturing which effectively removes unwanted outliers such as shadows from the operator and capturing equipment. A preference for a frontal view is desired as samples taken at grazing angles often exhibit artifacts from surface structure (visibility) and foreshortening effects. The spatial resolution term is chosen such that light probes captured closer to the surface are given high priority, as well as samples captured close to the center of the sphere in image coordinates.

To decrease visible seams between regions with different labels (pixels projected from different light probe images), we use a smoothness term $V_{p,q}(b_p, b_q)$ inspired by [39], where the goal is to reduce the difference in color and gradient across seams. We also enforce a penalty for using radiance samples l_c (from the light probes corresponding to the assigned label), with different spherical distortion across seams. The smoothness term is described by

$$V_{p,q}(b_p, b_q) = \begin{cases} 0 & \text{if } b_p = b_q \\ DC + DG + DSD & \text{otherwise} \end{cases} \quad (5)$$

and

$$\begin{aligned} DC &= |I_{b_p}(p) - I_{b_q}(p)| + |I_{b_p}(q) - I_{b_q}(q)| \\ DG &= |\nabla I_{b_p}(p) - \nabla I_{b_q}(p)| + |\nabla I_{b_p}(q) - \nabla I_{b_q}(q)| \\ DSD &= |D_{b_p}(p) - D_{b_q}(p)| + |D_{b_p}(q) - D_{b_q}(q)| \end{aligned} \quad (6)$$

where I represents the intensity of the sample, ∇I the gradient, and D the spherical distortion factor measured as the distance in image space from the center of the sphere to the sample.

To further smooth the possible seams left after the graph-cut optimization we use Poisson blending [40] to smooth the boundaries between patches of different labels (image probes). Fig. 7a shows an extracted proxy geometry surface corresponding to a brick wall textured by blending all captured radiance samples (representing the mean radiance), clearly blurring the texture. In Fig. 7b the brick wall is textured by solving the energy minimization problem using our proposed energy term.

6. Light field processing for lighting extrapolation

The generated light fields, L_A , are typically located outside of the region of space where light probes have been captured, Γ .

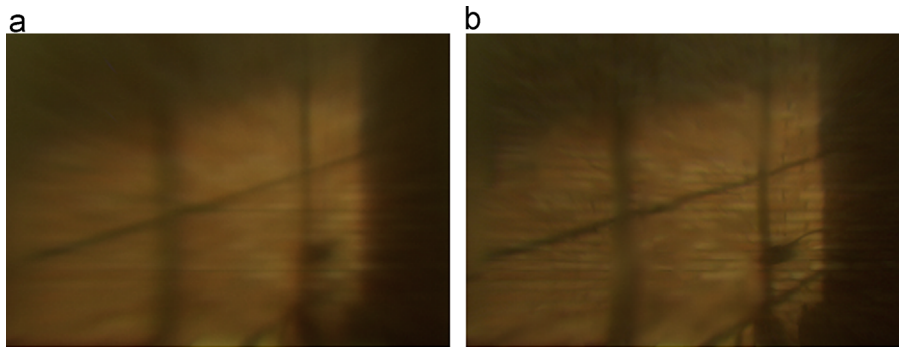


Fig. 7. Textured brick wall using (a) mean value of captured radiance samples (b) graph cut optimization considering geometric distortion of captured light probes.

As illustrated in Fig. 8a, this means that the re-projected radiance samples do not cover the entire hemisphere at each point on A_i . To ensure a predictable behavior when rendering objects outside Γ , the set of sampled angles at each spatial position on A_i needs to be extended to cover the entire hemisphere. For the extracted and re-projected light sources this is performed by filling in the missing values with a weighted average of the closest captured directions. If the sampling of a light source is not dense enough to accurately capture blockers in front of it, it is possible to interactively analyze the 4D data as described in [16] and enhance the resolution by placing one or more filters depicting high resolution blockers in the ray path. However, in many cases an even better result is achieved by removing the blocked radiance samples from the light field data set and introducing explicit 3D blocking geometry in the modeling step.

For surfaces, A_i , that coincide with geometry in the real scene, L_{A_i} will describe the reflected radiance of the surface. The scene model with re-projected radiance information makes it possible to estimate the radiance distribution incident at each point on A_i by sampling the extracted light sources and proxy geometries. Similar to [41] we use this information to recover a parametric material BRDF model by an optimization in the corresponding model parameter space. The user first selects a surface that is accurately described in the scene model. For each spatial point on the selected surface, we then compute the incident radiance distribution, and minimize the error between the measured data and the incident radiance transformed by the parametric BRDF. In our examples we have only considered the BRDF parameters yielding the smallest global error over all surface locations. The recovered BRDFs are then used either in their parametric form or to fill in the missing angular samples in L_{A_i} . This is performed by calculating the material response to the estimated incident radiance distribution at each point. For surfaces not modeled to match the scene geometry, BRDF recovery is not possible. Instead, missing values are filled in with weighted averages of the closest captured directions or the mean value, $\bar{w}(x, y, z)$, yielding an approximation sufficient for high quality IBL rendering. Fig. 8b and c displays renderings of a small subset of the light field surfaces from the example scene. Fig. 8b shows the unprocessed original 4D data with missing values. In Fig. 8c our 2D/4D data structure is used, with missing values filled in and a Blinn–Phong model fitted on the TV. Our approach makes it possible to render synthetic objects well outside Γ with high quality results.

7. Rendering

During rendering, the goal is to solve the rendering integral for spatially varying IBL in Eq. (3) for each sample point visible from the camera. At a point \mathbf{x} on a surface of the virtual objects, we want to efficiently sample the incident radiance distribution,

$\hat{L}_{SIBL}(\mathbf{x} \leftarrow \vec{\omega}_i)$, described by the recovered proxy geometry and the reprojected radiance along a set of directions $\vec{\omega}_i$, as illustrated in Fig. 9a. During rendering, this is carried out by sampling the recovered scene model using e.g. ray tracing or photon mapping. In most rendering packages this is in practice implemented so that the surfaces, A_i , in the model are represented as area or geometric light sources with a view dependent 2D/4D-textures and sampled proportional to their power.

To reconstruct the radiance contribution from a point \mathbf{P} observed from a surface point \mathbf{x} at an incident direction $\vec{\omega}_i$, we sample the underlying 2D/4D light field as illustrated in Fig. 9b. First, we determine the spatial position \mathbf{P} and its corresponding texture coordinates in uv -space and the direction $\vec{\omega}_i$ from \mathbf{P} to the point \mathbf{x} on the virtual object. The example in 9b illustrates two points with 2D information only, and two points where the full 4D distributions are stored. For the 2D case, the radiance contribution is considered to be diffuse and is fetched from the data structure. In the 4D case, we introduce a depth correction similar to [9,2], where we compute the direction from the surface point \mathbf{x} to the actual sample points on the A_i surface as illustrated by $\vec{\omega}_1$ and $\vec{\omega}_2$ in the figure. The radiance distribution at the spatial sample point is then sampled in the depth corrected direction. This is to increase rendering quality and remove ghosting artifacts. The final radiance contribution from the point \mathbf{P} incident at \mathbf{x} from the direction $\vec{\omega}_i$ is, in our implementations, reconstructed using bi-linear interpolation in both the spatial and angular domains.

8. Results

The software processing pipeline is implemented as a set of functions bundled into a library and interfaced through Autodesk Maya. For maximum performance the framework utilizes GPU processing where appropriate. Rendering is straightforward to implement in most rendering frameworks. The renderings presented here were done using V-Ray from Chaos Group, mental ray from NVIDIA, and the open source renderer PBRT [42].

To demonstrate the utility of HDR video for scene capture, image based lighting and photo realistic rendering, we show a number of reconstructed environments and how virtual objects can be placed and rendered to appear as if they were actually there. Rendering can be easily integrated into most modern renderers, and rendering times are well comparable to that of ordinary computer graphics rendering with purely virtual scenes.

As our first result, Fig. 1, we show a comparison of a rendering using our method to a traditional IBL rendering in which the lighting environment was captured as a single HDR light probe image in the same scene. It is evident that accurate representation of both spatial and angular variations in the scene lighting included in the scene model, Fig. 1b, clearly reaches higher levels of realism compared to the traditional IBL rendering, Fig. 1a. This is

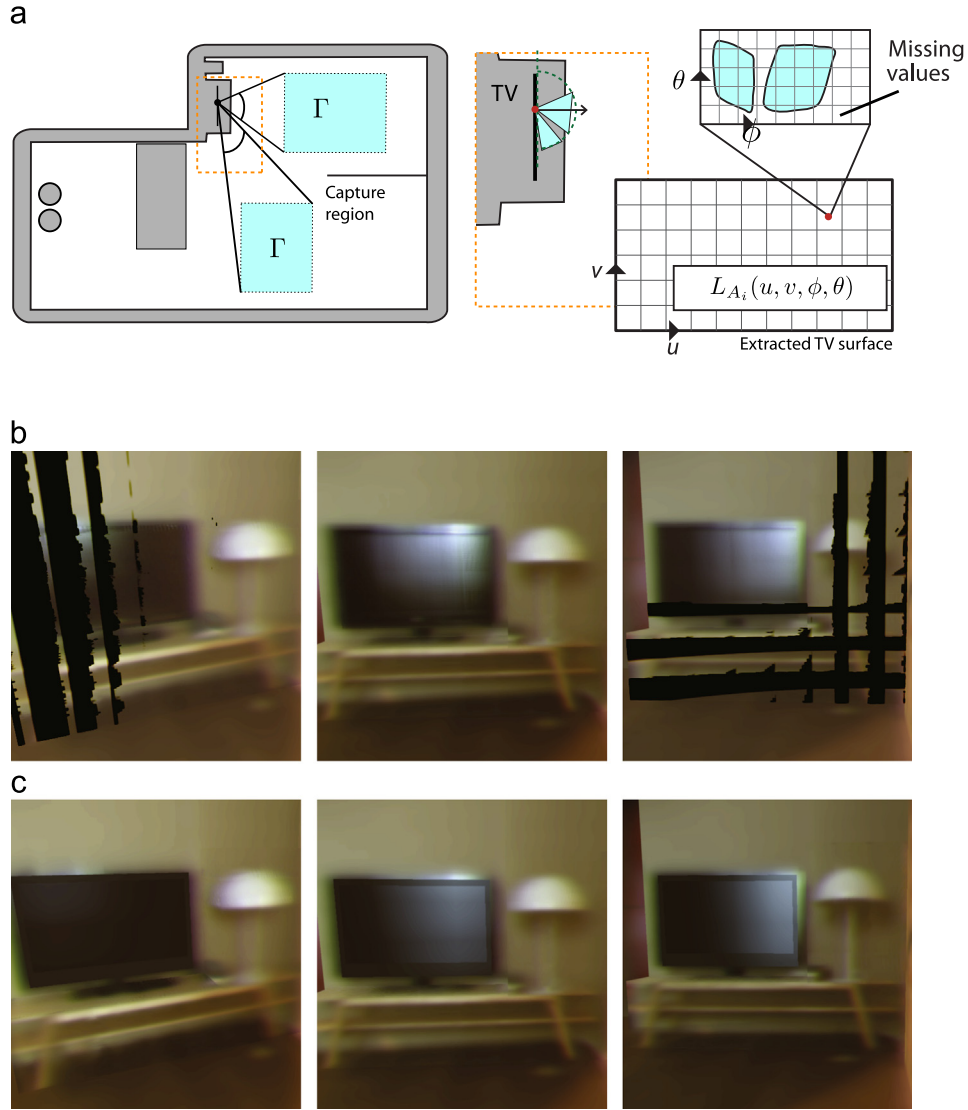


Fig. 8. (a) Illustrates how the finite size of the capture region Γ may lead to holes in the angular domain in the re-projected radiance distributions at the recovered surfaces A_i in the scene model. (b) Displays a set of renderings of the recovered scene model as the camera moves in and out of the capture region Γ . Rendering of 2D/4D surfaces before extrapolation (with holes) and (c) The holes can be filled in using the mean value as on the wall and lamp, or by fitting a parametric BRDF model to the measured data as on the TV. Rendering 2D/4D surfaces after extrapolation processing.

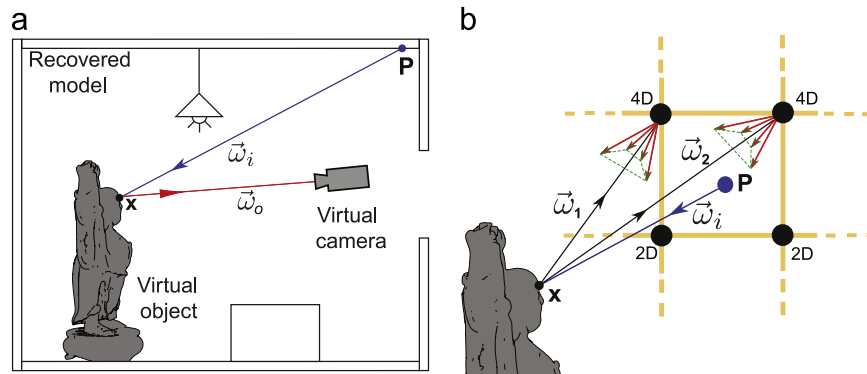


Fig. 9. Illustrates the rendering process. (a) The radiance contribution incident from a direction $\vec{\omega}_i$ at a point \mathbf{x} on a virtual object from a point \mathbf{P} at an A_i surface in the recovered model is (b) reconstructed by interpolation between neighboring radiance samples in the 2D/4D data structure stored at the surface.

also an example of how a scene can easily be remodeled, as it uses modeled illumination elements A_i (surfaces and light sources) from the photo studio displayed in Fig. 10a for lighting the synthetic floor, walls and furniture.

The second result is an overview of the full scene reconstruction, modeling and rendering of the photo studio displayed in Fig. 10a. The studio setup consists of three different example scenes: a kitchen area, a living room and a study. In each scene

two data sets, each covering a volume of $1.5 \times 1.5 \times 1.5$ m, were captured using a real-time light probe setup, see Fig. 3a, with a mirror sphere and SICK-IVP Ranger C55 cameras capable of capturing the full dynamic range in the scene. The setup was attached to a mechanical tracking stage, see Fig. 3d, measuring the xyz position of the light probe with sub-millimeter accuracy.



Fig. 10. (a) Shows the real photo studio that was captured and processed using the proposed pipeline. (b) displays a high resolution rendering of a virtual furnishing. The placement of virtual furnishes can be seen in Fig. 2 bottom.

Capture was performed with user guided adaptive sampling of the scene resulting in irregular spatial distribution of the light probe images within Γ . Each data set consists of around 30,000 HDR images of 512×512 pixels, corresponding to about 30 GB of data. This dense sampling of the scene is what guarantees capture of detailed spatial variations in the illumination, such as cast shadows and spotlights. For each scene, the \bar{w} and w_v volumes were computed at a resolution of $256 \times 256 \times 256$ voxels using a subset of 1500 light probes from each sequence. This resulted in a spatial resolution of about 2 cm per voxel. Running in 6 threads on a 2.8 GHz dual quad-core MacPro this took a few minutes per data set. Using the recovered point cloud, see Fig. 4a, as starting point, the proxy geometry of each scene was then modeled using planar surfaces aided by focal surface visualization, see Fig. 4b, and the light sources were extracted using the semi-automatic point and click interface as displayed in Fig. 5a. Before re-projection of the full set of radiance samples to the proxy geometries, the variance on each surface was visualized to let the user decide appropriate thresholds for selecting 2D or 4D storage of the light fields at each point as illustrated in Fig. 6. By automatic processes, the missing angles in the extracted light fields were extrapolated and a BRDF model was fitted to the specular surface of the TV set as shown in Fig. 8. Building proxy geometry at the level of detail displayed in the final scene model for the photo studio shown in Fig. 4b took approximately 1 h for an experienced Maya user. Most of that time was spent modeling the kitchen area. The living room area was modeled with less geometric detail and took only a few minutes. Virtual furniture was placed in the scene as illustrated in Fig. 2 (bottom). In Figs. 2 (bottom) and 10b, we display the level of realism attainable using our method. Fig. 10b shows a virtual furnishing against a backdrop of the original scene, and Fig. 2 a close-up view. It is evident that our technique yields improved levels of realism in the rendering of synthetic objects into real scenes. The images were rendered with a resolution of $4k \times 3k$ pixels in about an hour on a 2.93 GHz dual quad-core machine. This is similar to rendering times using ordinary computer graphics lights.

In our final example, shown in Fig. 11, we demonstrate the results from the capture, processing, and virtual furnishing of the living room and kitchen areas of a private summer house at an off-site location. An overview (latitude–longitude) panoramic image

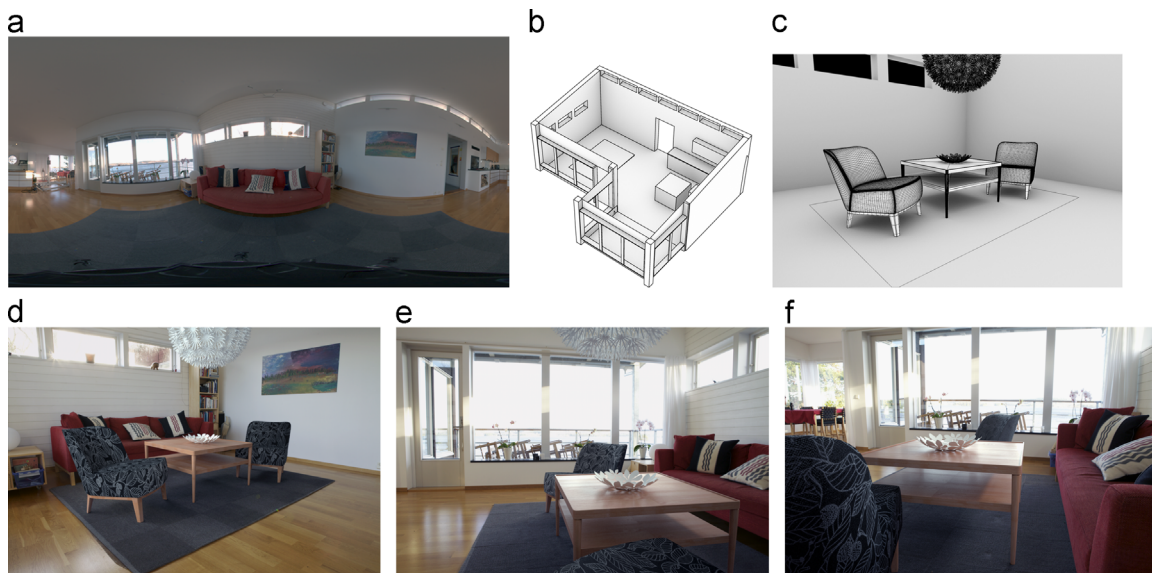


Fig. 11. (a) A 360° panorama showing an overview of the living room and kitchen areas of the captured summer house. (b) An ambient occlusion rendering displaying the geometric proxy model recovered from the scene. (c) Virtual furnitures were placed in the scene. (d), (e), and (f) Renderings generated with our framework along a camera path, showing how the real world lighting environment captured in the scene model interacts with the virtual objects.

of the room is displayed in Fig. 11a. The about 100 m² L-shaped room was captured using the multi-sensor HDR-video camera displayed in Fig. 3b and c as four light probe sequences: two in the living room area and two in the kitchen area. Each took in the order of 15 min. to capture, amounting to a total of 90,000–100,000 HDR light probe images. Including setup time, and time for capture of high resolution HDR video sequences for structure from motion computations and backdrops, the total time spent on site was about 2.5 h. This should be compared to real refurbishings and ordinary photo-shoots that often take 2–3 days (without the option of reshooting if necessary). The final proxy geometry, shown as a rendering in Fig. 11b, was recovered using structure from motion to get the point cloud data and then modified and modeled using the focal volume modeling approach. The total time to build and process the complete scene model, including re-projection of radiance samples, was in the order of 3–4 h for an experienced Maya user. Fig. 11c displays an example placement of virtual furnishings in the scene, and Fig. 11d, e and f is final renderings showing how the geometry and materials of the virtual furnitures interact with the captured lighting information and blend into the real scene.

9. Conclusion

This paper presented a set of novel algorithms and data structures that together with state-of-the-art HDR-video hardware and new capture methodologies form a complete spatially varying image based lighting pipeline for capture, processing and rendering virtual objects into real scenes with previously unattainable realism. The main contributions in this paper are the algorithms for robust and efficient scene reconstruction, extraction of light sources, radiance re-projection from spherical light probe images using a novel cost function for graph-cut optimization, a novel 2D/4D data structure for light field storage, and novel methods for lighting extrapolation outside the captured regions. Another key aspect of the paper is that it shows how HDR-video enables new applications in image based lighting and photo realistic rendering.

10. Limitations and future work

There are a number of interesting venues for future work. The technique described in this paper is currently limited to static scenes. In many cases this is a reasonable restriction since the time required (in our examples around 15–20 min.) also enables capture in less controlled outdoor environments. It is, however, of high interest to extend the processing with mechanisms for filtering out moving objects in the scene. The current implementation of the capture equipment limits the size of the photo sets to around 10 × 10 m. Portable capture equipment would allow for capture of larger sets and even more detailed light fields. Future work also includes investigation of methods for more efficient storage and representation of the light field data, as well as techniques for further assisting artists in the analysis of light sources and material properties.

Acknowledgments

This project was funded by the Swedish Foundation for Strategic Research (SSF) through Grant IIS11-0081, Linköping University Center for Industrial Information Technology (CENIIT), and the Swedish Research Council through the Linnaeus Environment CADICS.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version of <http://dx.doi.org/10.1016/j.cag.2013.07.001>.

References

- [1] Debevec P. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In: Proceedings of SIGGRAPH '98, 1998, p. 189–98.
- [2] Unger J, Wenger A, Gardner A, Hawkins T, Debevec P. Capturing and rendering with incident light fields. In: Proceedings of the 14th eurographics workshop on rendering, 2003, p. 141–9.
- [3] Masselus V, Peers P, Dutre P, Willems YD. Relighting with 4D incident light fields. ACM Transactions on Graphics 2003;22(3):613–20.
- [4] Unger J, Gustavson S. High-dynamic-range video for photometric measurement of illumination. In: Proceedings of sensors, cameras and systems for scientific/industrial applications X, IS&T/SPIE 19th international symposium on electronic imaging, vol. 6501, 2007.
- [5] Kronander J, Gustavson S, Bonnet G, Unger J. Unified HDR reconstruction from raw CFA data. In: Proceedings of the IEEE international conference on computational photography, 2013.
- [6] Kajiya JT. The rendering equation. In: SIGGRAPH computer graphics, ACM, vol. 4, 1986.
- [7] Reinhard E, Ward G, Pattanaik S, Debevec P. High dynamic range imaging—acquisition, display and image-based lighting. San Francisco, CA: Morgan Kaufmann; 2006.
- [8] Levoy M, Hanrahan P. Light field rendering. In: Proceedings of SIGGRAPH '96, 1996, p. 31–42.
- [9] Gortler SJ, Grzeszczuk R, Szeliski R, Cohen MF. The lumigraph. In: Proceedings of SIGGRAPH '96, 1996, p. 43–54.
- [10] Debevec P, Hawkins T, Tchou C, Duike HP, Sarokin W, Sagar M. Acquiring the reflectance field of a human face. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques, SIGGRAPH '00. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co.; 2000, p. 145–56.
- [11] Garg G, Talvala EV, Levoy M, Lensch H. Symmetric photography: exploiting data-sparseness in reflectance fields. In: Eurographics symposium on rendering (EGSR), 2006.
- [12] Adelson EH, Bergen JR. The Plenoptic function and the elements of early vision. In: Computational models of visual processing. Cambridge, MA: MIT Press; 1991 [chapter 1].
- [13] Sen P, Chen B, Garg G, Marschner SR, Horowitz M, Levoy M, et al. Dual photography. ACM Transactions on Graphics 2005;24(3):745–55.
- [14] Unger J. Incident light fields. PhD thesis, The Department of Science and Technology, Linköping University, The Institute of Technology; 2009. Linköping Studies in Science and Technology, Dissertations, No. 1233.
- [15] Ihrk I, Stich T, Gottschlich H, Magnor M, Seidel HP. Fast incident light field acquisition and rendering. Journal of WSCG 2008;16:25–32.
- [16] Unger J, Gustavson S, Larsson P, Ynnerman A. Free form incident light fields. Computer Graphics Forum 2008;27(4):1293–301.
- [17] Löw J, Ynnerman A, Larsson P, Unger J. HDR light probe sequence resampling for realtime incident light field rendering. In: Proceedings of SCCG'09, 2009, p. 53–60.
- [18] Karsch K, Hedau V, Forsyth, D, Hoiem, D. Rendering synthetic objects into legacy photographs. In: Proceedings of the 2011 SIGGRAPH Asia conference, SA '11. New York, NY, USA: ACM; 2011, p. 157:1–12.
- [19] Khan EA, Reinhard E, Fleming RW, Bühlhoff HH. Image-based material editing. ACM Transactions on Graphics 2006;25(3):654–63.
- [20] Sinha SN, Steedly D, Szeliski R. Piecewise planar stereo for image-based rendering. In: 2009 IEEE 12th international conference on computer vision, 2009, p. 1881–8.
- [21] Furukawa Y, Curless B, Seitz SM, Szeliski R. Reconstructing building interiors from images. In: Proceedings of the IEEE international conference on computer vision, 2009, p. 80–7.
- [22] Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: Proceedings of CVPR '06, 2006, p. 519–28.
- [23] Newcombe RA, Davison AJ, Izadi S, Kohli P, Hilliges O, Shotton J, et al. KinectFusion: real-time dense surface mapping and tracking. In: 2011 10th IEEE international symposium on mixed and augmented reality (ISMAR), IEEE; 2011, p. 127–36.
- [24] Gruber L, Richter-Trummer T, Schmalstieg D. Real-time photometric registration from arbitrary geometry. In: Proceedings of the 2012 IEEE international symposium on mixed and augmented reality (ISMAR), ISMAR '12; Washington, DC, USA: IEEE Computer Society; 2012, p. 119–28.
- [25] Debevec PE, Taylor CJ, Malik J. Modeling and rendering architecture from photographs: a hybrid geometry- and image-based approach. In: Proceedings of SIGGRAPH '96, p. 11–20. ISBN 0-89791-746-4, 1996.
- [26] van den Hengel A, Dick A, Thormählen T, Ward B, Torr PHS. Videotrace: rapid interactive scene modeling from video. In: Proceedings of SIGGRAPH '07, ACM; 2007, p. 86–96.

- [27] Sinha SN, Steedly D, Szeliski R, Agrawala M, Pollefeys M. Interactive 3D architectural modeling from unordered photo collections. *ACM Transactions on Graphics* 2008;27(5):159:1–10.
- [28] Isaksen A, McMillan L, Gortler SJ. Dynamically reparameterized light fields. In: *Proceedings of SIGGRAPH '00*, 2000. p. 297–306.
- [29] Vaish V, Levoy M, Szeliski R, Zitnick CL, Kang SB. Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures. In: *Proceedings of CVPR '06*, 2006. p. 2331–8.
- [30] Unger J, Gustavson S, Ollila M, Johannesson M. A real time light probe. In: *Proceedings of the 25th eurographics annual conference, vol. short papers and interactive demos*, 2004. p. 17–21.
- [31] Tocci MD, Kiser C, Tocci N, Sen P. A versatile HDR video production system. *ACM Transactions on Graphics* 2011;30(4):41:1–10.
- [32] Snavely N, Seitz SM, Szeliski R. Photo tourism: exploring photo collections in 3D. *ACM SIGGRAPH 2006 Papers. SIGGRAPH '06*. Boston, MA: ACM; 2006;835–46. ISBN: 1-59593-364-6.
- [33] Furukawa Y. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2010;32(8):1362–76.
- [34] Lempitsky V, Ivanov D. Seamless mosaicing of image-based texture maps. In: *IEEE conference on computer vision and pattern recognition*, 2007, CVPR '07, 2007. p. 1–6.
- [35] Gal R, Wexler Y, Ofek E, Hoppe H, Cohen-Or D. Seamless montage for texturing models. *Computer Graphics Forum* 2010;29(2):479–86.
- [36] Boykov Y, Veksler O, Zabih R. Efficient approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2001;20(12):1222–39.
- [37] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2004;26:359–74.
- [38] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2004;26(2):147–59.
- [39] Agarwala A, Dontcheva M, Agrawala M, Drucker S, Colburn A, Curless B, et al. Interactive digital photomontage. *ACM Transactions on Graphics* 2004;23(3):294–302.
- [40] Pérez P, Gangnet M, Blake A. Poisson image editing. *ACM Transactions on Graphics* 2003;22(3):313–8.
- [41] Yu Y, Debevec P, Malik J, Hawkins T. Inverse global illumination: recovering reflectance models of real scenes from photographs. In: *Proceedings of SIGGRAPH '99*, 1999. p. 215–24.
- [42] Pharr M, Humphreys G. *Physically based rendering: from theory to implementation*. 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2010.